

Научная статья
УДК 551.515.4: 528.85
doi: 10.17223/25421379/24/11

ВЫДЕЛЕНИЕ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ ВЕРОЯТНОСТНЫХ ЗОН РАЗВИТИЯ ГРОЗЫ ПО ДАННЫМ СПУТНИКОВОГО ЗОНДИРОВАНИЯ



Владислав Вячеславович Чурсин¹, Ирина Валерьевна Кужевская²

^{1,2} Национальный исследовательский Томский государственный университет, Томск, Россия

¹ skriptym@mail.ru

² irina-kuz@vtomske.ru

Аннотация. Рассмотрена возможность детектирования зон грозы на основе данных микроволновых измерений приборов AMSU/MHS и ATMS, а именно вертикальных профилей метеопараметров (температуры и удельного влагосодержания), восстановленных в специализированном программном комплексе MIRS. Расчеты осуществляются с применением разработанной и обученной авторами нейронной сети прямого распространения. Для территории Сибири отмечаются значения общей оправдываемости детектирования гроз более 84 %.

Ключевые слова: опасные явления, реанализ ERA5, детектирование гроз, конвективная облачность

Источник финансирования: исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-35-90059.

Для цитирования: Чурсин В.В., Кужевская И.В. Выделение с помощью нейронных сетей вероятностных зон развития грозы по данным спутникового зондирования // Геосферные исследования. 2022. № 3. С. 162–171. doi: 10.17223/25421379/24/11

Original article
doi: 10.17223/25421379/24/11

NEURAL NETWORK DETECTION OF POTENTIAL THUNDERSTORM ZONES FROM REMOTE SENSING DATA

Vladislav V. Chursin¹, Irina V. Kuzhevskaya²

^{1,2} National Research Tomsk State University, Tomsk, Russia

¹ skriptym@mail.ru

² irina-kuz@vtomske.ru

Abstract. Convection origination on the territory of Eurasia is noted up to 65–67° north latitudes, as noted in the annual reports on thunderstorm registration by Vaisala. Registration of thunderstorms by geostationary satellites in this territory is difficult. Therefore, the question of using satellite data from polar-orbiting spacecraft to monitor zones of convection development is relevant.

The objectives of this study were to develop an architecture of algorithms for detection of atmospheric phenomena based on machine learning and neural networks, creation and training of algorithms for thunderstorm detection using ERA5 reanalysis data and further verification of the obtained algorithms and models on satellite sensing data of the MIRS program complex.

The technology of selecting probable zones of thunderstorm development from satellite sensing data is a model of probabilistic detection of the presence or absence of atmospheric phenomena. With the help of calculations, the zones in which the atmospheric parameters more or less correspond to the conditions under which the hydrometeorological phenomena can form are highlighted. Calculations are carried out with the help of machine learning technology and neural networks.

The paper presents the architecture of an algorithm for thunderstorm detection based on machine learning and neural network technology. A fully connected neural network is used, where the output signal of each neuron is fed as an input signal to all subsequent neurons. The neural network includes forty input neurons, twenty neurons on the first hidden layer, fifteen neurons on the second hidden layer, and two neurons corresponding to the necessary classification – presence/absence of thunderstorm on the output layer. The activation function on the hidden layers is Rectified linear unit (ReLU), the activation function on the output layer is Softmax. The algorithm was trained on the data of reanalysis ERA5 and verified on the data of satellite sounding of program complex MIRS which uses microwave measurements of AMSU/MHS devices of NOAA series, MetOp and ATMS device of Suomi NPP. Based on the results of the analysis of a number of data that were not involved in the model development and training process for May–October 2019–2020, an accuracy of 84 % was obtained.

Additionally, tests were conducted on 109 meteorological stations located within the boundaries of 49.9°–60.35° N and 75.68°–88.67° E. For the periods: April–September 2021, the predictability value indicates that 96 % of the events are successfully classified (presence/absence of thunderstorms).

The presented algorithm is dynamic. The learning procedure can be reinitialized when enough data is accumulated. This will allow us to take into account the appearance of new extremes of atmospheric characteristics.

The use of the presented algorithm and the results of its calculations is promising as an additional tool for operational work.

Keywords: hazard phenomena, ERA5 reanalysis, thunderstorm detection, convective cloudiness

Source of financing: The reported study was funded by RFBR, project number 20-35-90059.

For citation: Chursin V.V., Kuzhevskaya I.V. (2022) Neural network detection of potential thunderstorm zones from remote sensing data. *Geosfernye issledovaniya – Geosphere Research*. 3. pp. 162–171. (In Russian). doi: 10.17223/25421379/24/11

Введение

Важность оперативного определения зон грозы связана с комплексностью этого явления, причем представляют угрозу не только непосредственно молниевая опасность, но и сильный ветер и сильные осадки во время прохождения грозовой облачности.

Корректное воспроизведение профилей температуры и влагосодержания алгоритмами обработки спутниковых спектральных данных определяет качество детектирования зон грозовой деятельности.

Непосредственно краткосрочные прогнозы гроз основаны, как правило, на оценке возможности развития запасов энергии конвективной неустойчивости атмосферы. Подобная оценка стратификации и влагосодержания проводится на основе данных моделей численных прогнозов погоды (ЧПП) различного масштаба. На настоящем этапе для прогноза ЧПП остается нерешенной проблема внутримассовых гроз [Иванова, 2019], относящаяся к вопросам подсчетного масштаба и краткосрочности самого явления. Особенно серьезную проблему представляют мультячейковые или суперячейковые конвективные штормы с размером ячеек от 10 до 100–150 км, быстро развивающиеся по высоте, наблюдения за которыми с помощью радиолокатора описаны в работе [Абдулаев, Ленская, Желнин, 2012]. Подобные явления не прогнозируются и относятся к ряду опасных явлений.

Как отражено в работах [Иванова, 2019; Meng, Yao, Xu, 2019; Lee et al., 2020], именно с задачей быстрого распознавания зон грозы и произошло зарождения понятия «наукастинг». По мере развития спутниковой метеорологии появилась возможность выпускать предупреждения о грозовой облачности, в первую очередь, на основе информации с геостационарных спутников GOES и Meteosat. Достаточно много опубликовано работ, посвященных трекингу Cb облачности, например [Galanaki et al., 2018], по данным геостационарных спутников.

Известно, что на спутниковых данных видимого и ИК-диапазонов *Cu cong* и *Cb* в районах с грозами имеют ряд тональных особенностей и количествен-

ных значений параметров, характерных для этого атмосферного явления. Авторы [Алексеева и др., 2006] развивали подход к распознаванию осадков любого типа, гроз и оценке их интенсивности по снимкам ИК-диапазона, получаемым с геостационарных искусственных спутников Земли. В работе [Алексеева, Бухаров, 2005] представлена методика автоматизированного распознавания гроз по информации с космического аппарата (КА) серии NOAA, которая показывала хорошую согласованность с данными наземной сети. Оценка вероятности возникновения гроз по данным спутниковых измерений в ИК-диапазоне и синхронной прогностической метеорологической информации приведена в работе [Бухаров, 2013], в которой приводится анализ соответствия между результатами пеленгации гроз разными сетями.

В случае спутниковых и радиолокационных данных, соединение скоплений молний в трек может быть сложной задачей, поскольку в одном районе может быть несколько грозовых ячеек или несколько гроз (которые также могут разделяться и сливаться) [del Moral, Rigo, Llasat, 2018]. Непрерывный мониторинг молний над Северной и Южной Америкой, Тихим и Атлантическим океанами обеспечивается данными с геостационарных спутников программы GOES, которые были запущены в 2016 и 2017 гг. [Goodman et al., 2012]. Himawari-8 расширяет зону мониторинга Тихого и добавляет зону Индийского океана. В целом, все геостационарные метеорологические спутники обладают уникальными преимуществами в мониторинге характеристик начальных стадий конвективных облаков, таких как быстрое увеличение мощности облака, резкое снижение температуры и фазовая трансформация в вершине облака [Mecikalski et al., 2010]. Поэтому спутниковые данные могут быть использованы для мониторинга зарождения конвекции и ее раннего предупреждения для территорий до 40–45° широты.

Зарождение конвекции на территории Евразии отмечается и в более северных широтах до 65–67° [Network V.L.D., 2021], более того, в последние десятилетия отмечается смещение границ активной

конвекции на север [Горбатенко и др., 2020], а по данным [Чернокульский и др., 2018], на территории исследования в целом доминируют осадки ливневого и смешанного характера. В связи с чем актуален вопрос использования спутниковых данных с полярно-орбитальных космических аппаратов для мониторинга зон развития конвекции севернее 45° широты.

В задачу настоящего исследования входила разработка архитектуры алгоритма детектирования атмосферных явлений. Алгоритм основан на технологиях машинного обучения и нейронных сетях, его создание и обучение проходили на данных реанализа ERA5, верификация полученных алгоритмов и моделей – на данных спутникового зондирования программного комплекса (ПК) MIRS.

Материалы и методы исследования

Технология выделения вероятных зон развития гроз по данным спутникового зондирования представляет собой модель вероятностного детектирования наличия или отсутствия атмосферного явления. С помощью расчетов выделяются зоны, в которых параметры атмосферы в большей или меньшей степени соответствуют условиям, при которых может сформироваться атмосферное явление. Расчеты осуществляются с помощью технологии машинного обучения и нейронных сетей.

Наземные данные. За эталонные данные о фактическом наблюдении грозы на станции, на этапе валидации нейронной сети, была принята информация из штормовых телеграмм в коде WAREP, предоставленная ФГБУ «Западно-Сибирское УГМС», за период май–октябрь 2021 г. Обучение нейронной сети и тестирование проводились на основе информации из открытой базы данных ВНИИГМИ МЦД [Булыгина и др.], содержащей данные о типах атмосферных явлений, их интенсивности и времени начала/окончания, за май – октябрь 1990–2020 гг. по 135 метеорологическим станциям, расположенным в границах 50° – 70° с.ш. и 50° – 100° в.д. Исследование выполнено на оборудовании Уникальной научной установки «Система экспериментальных баз, расположенных вдоль широтного градиента» ТГУ при финансовой поддержке Минобрнауки России (RF-2296.61321X0043, 13.УНУ.21.0005, договор № 075-15-2021-672).

Данные реанализа ERA5. Реанализ ERA5 является продуктом ECMWF и содержит почасовые данные о параметрах поверхности суши, моря и атмосферы (для 37 уровней давления, от подстилающей поверхности до 80 км) с 1979 года по настоящее время. Пространственное разрешение – $0,25^\circ \times 0,25^\circ$. ERA5 является результатом ассимиляции данных 4D-Var в

CY41R2 интегрированной прогнозной системе ECMWF (IFS), которая усваивает данные наземной наблюдательной сети и данные аэрологического зондирования [Hersbach et al., 2020]. Из всех данных реанализа использовались данные о вертикальном распределении температуры и влажности. На их основе рассчитывались вертикальные профили температуры точки росы, относительной влажности, дефицита точки росы и значения индексов неустойчивости (Vertical Totals, Cross Totals, Total Totals, K-индекс и высоты нижней границы конвективной облачности).

Данные наземных наблюдений (135 метеорологических станций) и реанализа были совмещены в пространстве и времени. Для этого использовалась разность во времени наблюдений не более 5 мин, профили метеопараметров строились по ближайшему к метеорологической станции узлу сетки. Использовались данные за май–октябрь 1990–2018 гг. Полная выборка данных включала в себя 325 245 случаев.

Данные ПК MIRS. ПК MIRS разработан Центром спутниковых приложений и исследований NOAA/NESDIS (STAR) [User Manual, 2016]. Входными данными, используемыми для обработки в ПК MIRS, являются микроволновые измерения приборов AMSU/MHS KA серии NOAA, MetOp и прибора ATMS KA Suomi NPP. Для расчета параметров атмосферы и подстилающей поверхности ПК MIRS использует физически обоснованную модель восстановления метеорологических характеристик 1DVAR [Algorithm..., 2006]. Продукты программного комплекса зарекомендовали себя в мировой практике в качестве достоверного источника данных о распределении метеопараметров по высотам [Grassotti et al., 2020]. Точности восстановления метеопараметров соответствуют рекомендованным Всемирной метеорологической организацией, разработчиками проводится ежедневная оценка качества данных.

Данные наземных наблюдений (135 метеорологических станций) и спутникового зондирования были совмещены в пространстве и времени. Для этого использовалась разность во времени наблюдений не более 15 мин, профили метеопараметров строились по ближайшей к метеорологической станции точке зондирования. Использовались данные за май–октябрь 2019–2020 гг. Полная выборка данных включала в себя 1 933 случая.

Нейронная сеть. По структуре связей нейронных сетей в работе использована полносвязная нейронная сеть, где выходной сигнал каждого нейрона подается в качестве входного сигнала всем последующим нейронам. Выходными сигналами сети могут быть все или некоторые выходные сигналы нейронов. В исследовании использовалась открытая про-

граммная библиотека для машинного обучения, разработанная компанией Google-TensorFlow [Орельен, 2018], дополненная нейросетевой библиотекой KERAS [Джулли, Пал, 2017].

Модель нейронной сети принимает на вход значения параметров атмосферы и рассчитанных индексов неустойчивости. Архитектура разработанной авторами нейронной сети выглядит следующим образом: 40 входных нейронов, функция активации на скрытых слоях – Rectified linear unit (ReLU), функция активации на выходном слое – Softmax.

После попадания параметров на слой входных нейронов случайным образом отключаются 20 % нейронов для того, чтобы избежать переобучение модели. Далее данные поступают на еще один слой с функции активации ReLU, но с меньшим количеством нейронов – 20. После этого происходит нормализация выхода предыдущих слоев, после чего также отключаются 20 % случайных нейронов в слое. Затем данные поступают на еще один скрытый слой, содержащий 15 нейронов, после чего происходит передача уже модифицированных данных на

выходной слой, содержащий два нейрона, соответствующих необходимой классификации – наличие/отсутствие грозы. На выходном слое функцией активации является Softmax. Это позволяет на выходе получить процент вероятности соответствия метеорологических параметров тем, при которых наблюдаются грозовые явления. Функцией оптимизации для нейронной сети выступает NAdam, так как при применении функции Adam результаты были несколько хуже. Коэффициент скорости обучения (learning rate) нейронной сети составил 0,003 на обучении.

Метрики точности. В соответствии с Методическими указаниями по проведению производственных (оперативных) испытаний новых и усовершенствованных методов гидрометеорологических и гелиогеофизических прогнозов [РД 52.27.284-91, 1991], для установления надежности расчетных методов прогноза такого метеорологического явления, как гроза, первоначально составляется таблица сопряженности прогноза с фактическими данными о явлениях по форме 9, пример которой представлен в табл. 1.

Таблица 1

Таблица сопряженности прогноза с фактическими данными о явлениях (Форма 9)

Table 1

Table of contingency forecast with actual data on phenomena (Form 9)

Прогноз	Наблюдалось		Сумма	U	U + П
	Явление	Без явления			
Явление	n11	n12	n10	Uя	Uя + Пя
Без явления	n21	n22	n20	Uб.я	Uб.я + Пб.я
Сумма	n01	n02	n00	–	–
П	Пя	Пб.я	–	U	–
			T	H	

На основании этих данных рассчитываются оправдываемости в процентах прогнозов наличия (Uя) и отсутствия (Uб.я) явления. Для характеристики успешности метода прогнозов с учетом случайных прогнозов рассчитывается критерий надежности Н.А. Багрова (Н). Установлено, что прогнозы с $H < 0,33$ являются ненадежными. Для оценки качества альтернативных прогнозов также применен критерий качества прогнозов Пирси–Обухова (Т). Положительные значения Т свидетельствуют о том, что предупрежденность явления (Пя) превышает отношение ошибочных прогнозов явления к фактическому числу дней без явления. При идеальном прогнозе $T = 1$.

Дополнительные оценки качества:

– U общая оправдываемость, %;

– (Uя + Пя) суммарный показатель, характеризующий удовлетворительное качество прогнозов при условии $(Uя + Пя) > 130 \%$.

Обучение нейронной сети

Обучение нейронной сети – это процесс, при котором происходит настройка параметров нейронной сети посредством моделирования среды, в которую эта сеть встроена. В процессе обучения нейронная сеть выявляет сложные зависимости между входными и выходными данными и выполняет обобщение.

Для обучения нейронной сети применялся метод обучения с учителем. Процесс обучения с учителем представляет собой предъявление нейронной сети

выборки эталонных обучающих примеров. За эталонные данные были приняты данные наземной наблюдательной сети и данные реанализа ERA5. Полная выборка включала в себя 1 325 245 случаев за май – октябрь 1990–2018 гг. и содержала данные о метеорологических явлениях и сопутствующих им значениях характеристик атмосферы.

Весь массив данных был случайным образом разделен на обучающую и тестовую выборки в соотношении 70 % (927 672 случая) на 30 % (397 573 слу-

чая) с сохранением количественного распределения явлений. На выборке, состоящей из 927 672 случаев, проводилось обучение, а на выборке, состоящей из 397 573 случаев, – проверка. Схематично процедура обучения представлена на рис. 1.

Обучение модели производилось путем циклического подбора весовых коэффициентов до момента, пока значения точности (общей оправдываемости) на тестовой выборке не перестают увеличиваться на протяжении 10 последовательных эпох.



Рис. 1. Схема процедуры обучения

Fig. 1. Scheme of the training procedure

Согласно рис. 2, на первых эпохах точность нейронной сети быстро возрастает, но затем выходит на плато. Благодаря низкому значению коэффициента скорости обучения получается определить, что плато является ложным и точность на нем не максимальная. Наилучшей точности на тестовой выборке, при описанной архитектуре, удалось достигнуть к 174-й эпохе обучения со значением 0,76 или 76 %.

После обучения на данных реанализа ERA5, достоверность модели оценивалась на спутниковых данных ПК MIRS за период, не участвовавший в обучении (май–октябрь 2019–2020 гг.). При оценке достоверности выделения гроз принимался порог в

50 % – при достижении этого значения вероятности считается, что гроза фиксируется. Оценка производилась в строгом соответствии с наблюдениями на станции. В табл. 2 представлены результаты выделения зон гроз в соответствии с РД 52.27.284-91.

Как видно из табл. 2, общая оправдываемость составляет 84,3 %. Оправдываемость наличия грозы составляет 84,0 %, отсутствия – 84,7 %. Критерий надежности Н.А. Багрова составляет 0,68, что больше порогового значения 0,33 и говорит о надежности прогноза. Критерий качества прогнозов Пирси-Обухова составляет 0,69, т.е. предупрежденность явления превышает отношение ошибочных прогнозов явления к фактическому числу дней без явления. Суммарный показатель

(Уя + Пя) составляет 168,9 % и характеризует качество прогноза как удовлетворительное.

Таким образом, результаты оценки обучения нейронной сети для детектирования вероятных зон развития гроз над территорией Сибири с использованием открытой базы данных ВНИИГМИ МЦД

показали, что разработанная архитектура нейронной сети и описанный выше процесс обучения способствуют успешному детектированию гроз как на данных реанализа ERA5 (общая оправдываемость 76 %), так и на данных спутникового зондирования ПК MIRS (общая оправдываемость 84 %).

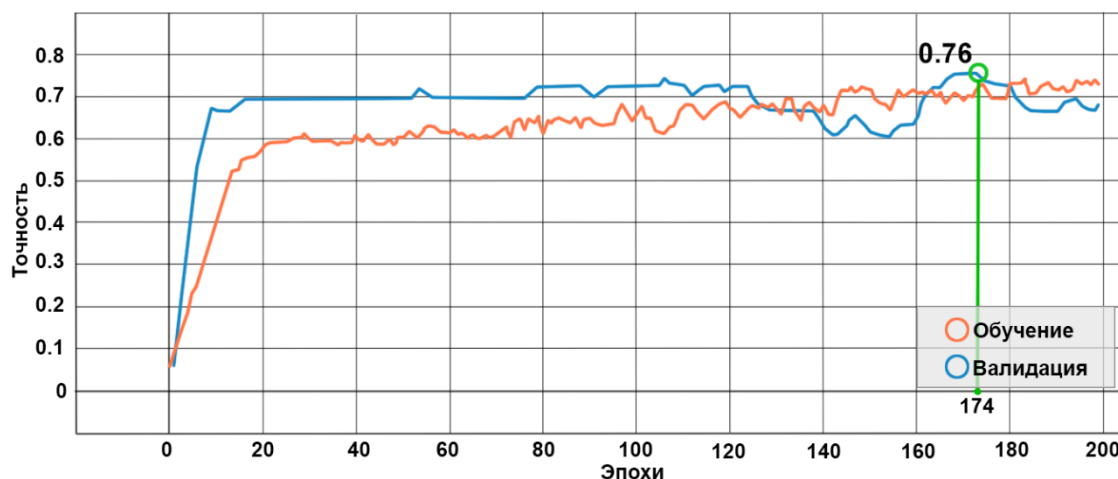


Рис. 2. Точность нейронной сети в зависимости от количества пройденных эпох

Fig. 2. Neural network accuracy as a function of the number of epochs traversed

Таблица 2

Таблица сопряженности выделения гроз методикой с фактическими данными о грозах на станциях

Table 2

Contingency table of detection of thunderstorms by a neural network with actual data on thunderstorms at stations

Прогноз	Наблюдалось		Сумма	U	U + П
	Явление	Без явления			
Явление	821	156	977	84,03	168,93
Без явления	146	810	956	84,72	168,49
Сумма	967	966	1933	84,30	
П	84,90	83,76	0,69	0,68	

Результаты испытания в оперативном режиме

С целью оценки достоверности детектирования гроз и определения возможности использования данной технологии в подразделениях Росгидромета проведены дополнительные испытания в период с 1 апреля по 30 сентября 2021 г.

Эталонными данными для сравнения являлись данные фактических наблюдений по станциям региона в коде КН-01 SYNOP и штормовых телеграмм кода WAREP. Оценке качества подлежал факт нали-

чия грозы в районе метеостанции с вероятностью более 80 %.

Технология испытывалась для территорий Новосибирской, Кемеровской, Томской областей, Алтайского края и Республики Алтай. Оценке подвергались результаты расчетов технологии выделения вероятных зон гроз. Оценка достоверности расчетов методики проводилась в соответствии с руководящим документом РД 52.27.284-91.

Для проведения испытаний данные наземных наблюдений проходили процедуру пространственного

и временного совмещения с данными дистанционного зондирования. Для временного совмещения задавалась минимально возможная разница (не более 10 мин) во времени между пролетом космического аппарата и наблюдениями на станции. Пространственное совмещение заключалось в восстановлении вертикальных профилей по спутниковым измерениям непосредственно над метеорологической станцией.

При испытаниях привлекались данные наблюдений на 109 метеорологических станциях ФГБУ «Западно-Сибирское УГМС», расположенных в границах 49,9–60,35° с.ш. и 75,68–88,67° в.д. За периоды испытаний (апрель–сентябрь 2021 г.) грозовые явления отмечались на всех 109 станциях.

Для оценки достоверности детектирования вероятных зон гроз использовалась выборка данных метеостанций, содержащая в себе:

- индекс станции;
- дату;
- данные раздела 1 группы 7 кода КН-01 об погоде в срок наблюдения и прошедшей погоде.

Итоговый массив содержал 1 088 случаев, из которых 76 случаев наличия грозового явления на станции в момент пролета КА и 1 012 с отсутствием грозового явления. В процентном соотношении число случаев с грозами и без гроз составляло 7 и 93 % соответственно.

Оценка достоверности производилась для всей территории за каждый месяц и за весь период испытаний для субъектов Российской Федерации, нахо-

дящихся в зоне ответственности ФГБУ «Западно-Сибирское УГМС». В табл. 3 представлены матрицы сопряженности и статистические характеристики оправдываемости детектирования гроз с вероятностью >80 % над станцией для всей территории по месяцам, а также обобщены результаты за период испытаний с разделением по субъектам.

При детектировании гроз над станцией в ежемесячной оценке максимальное значение общей оправдываемости (99,4 %) и максимальное значение Критерия Пирси–Обухова (0,99) отмечаются в сентябре, однако стоит учесть, что в этот месяц, в синхронные с пролетом космического аппарата сроки, была отмечена всего одна гроза. Наибольшее количество гроз, попавших в синхронные сроки, приходится на июль. Значение общей оправдываемости составляет 92,8 %. При этом значение критерия Пирси–Обухова равен 0,6, а критерий надежности по Н.А. Багрову – 0,62, что превышает пороговое значение 0,33 и говорит о надежности детектирования.

В целом замечено, что вне зависимости от месяца технологией успешно детектируется большая часть случаев с грозовыми явлениями. Среднее значение предупрежденности за весь период испытаний составляет 72,6 %, оправдываемость наличия явления 77,6 %. Значения критериев Пирси–Обухова, Н.А. Багрова и суммарный показатель ($U_{\text{я}} + P_{\text{я}}$) говорят о надежности детектирования. Значение общей оправдываемости свидетельствует о том, что успешно классифицируется в среднем 95,9 % событий.

Таблица 3

Матрицы сопряженности и статистические характеристики оправдываемости детектирования гроз с вероятностью >80 % над станцией для всей территории и по субъектам

Table 3

Contingency matrices and statistical performance characteristics of detection of thunderstorms with a probability of >80 % over the station for the entire territory and by subjects

Параметр	n11	n12	n21	n22	nn	$U_{\text{я}}$	$P_{\text{я}}$	U	T	H	$U_{\text{я}} + P_{\text{я}}$
Апрель	–	–	–	–	–	–	–	–	–	–	–
Май	12	0	3	214	229	100	80	98,7	0,8	0,9	180
Июнь	6	1	3	53	63	85,7	66,6	93,7	0,6	0,7	152,3
Июль	18	8	10	215	251	69,2	64,28	92,8	0,6	0,6	133,5
Август	12	0	11	197	220	100	52,2	95	0,5	0,7	152,1
Сентябрь	1	2	0	322	325	33,3	100	99,4	0,9	0,5	133,3
За период испытаний	49	11	27	1 001	1 088	77,6	72,6	95,9	0,7	0,7	150,2
Томская область	19	0	5	355	359	100	79,2	98,6	0,5	0,6	179,1
Новосибирская область	8	2	8	310	328	80	50	96,9	0,5	0,6	130
Кемеровская область	6	2	5	225	238	75	54,5	97	0,5	0,6	129,5
Алтайский край	11	6	8	155	180	64,7	57,9	92,2	0,5	0,6	122,6
Республика Алтай	5	2	2	82	91	71,42	71,4	95,6	0,7	0,7	142,8

При детектировании гроз над станцией в зависимости от субъекта замечено, что наибольшее значение общей оправдываемости (98,6 %) имеют станции, расположенные на территории Томской области. Немного хуже грозы детектируются над станциями Алтайского края и Республики Алтай, но при этом значение общей оправдываемости не опускается ниже 92 %. Наибольшую оправдываемость наличия явления также имеет Томская область (100 % гроз успешно детектированы). Значения критериев Пирси–Обухова и Н.А. Багрова говорят об удовлетворительном качестве детектирования на территории верификации.

Дополнительно были проанализированы случаи ложного детектирования гроз над станцией с целью

выявления атмосферных явлений, при которых они происходили. Распределение атмосферных явлений при случаях ложного детектирования гроз над станцией представлено на рис. 3.

Как можно заметить, ложное детектирование гроз происходило преимущественно при наблюдении на станции явлений, связанных с развитием конвекции в атмосфере, например ливневых осадков слабой интенсивности. Ложного детектирования при полном отсутствии атмосферных явлений, связанных с конвекцией, на станции не наблюдалось. Также стоит отметить, что при грозах между сроками наблюдений в 18 % случаев технология детектировала грозы.

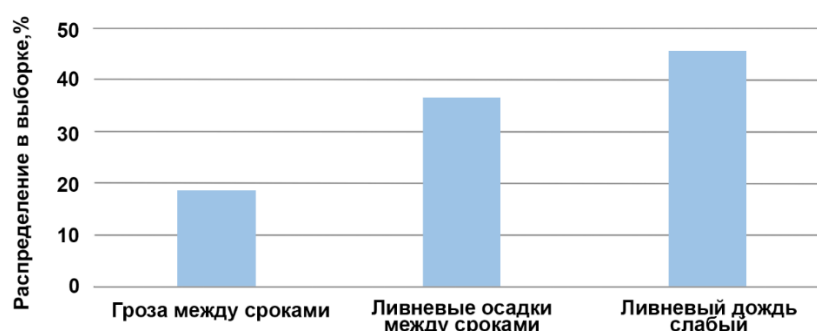


Рис. 3. Распределение атмосферных явлений при случаях ложного детектирования гроз над станцией

Fig. 3. Distribution of atmospheric phenomena in cases of false detection of thunderstorms over the station

Заключение

В работе приводится разработанная архитектура алгоритма детектирования гроз, основанного на технологиях машинного обучения и нейронных сетях. Использована полносвязная нейронная сеть, где выходной сигнал каждого нейрона подается в качестве входного сигнала всем последующим нейронам.

Нейронная сеть включает в себя 40 входных нейронов, 20 нейронов на первом скрытом слое, 15 нейронов на втором скрытом слое и 2 нейрона, соответствующих необходимой классификации, – наличие/отсутствие грозы на выходном слое. Функция активации на скрытых слоях – Rectified linear unit (ReLU), функция активации на выходном слое – Softmax. Алгоритм обучен на данных реанализа ERA5 и верифицирован на данных спутникового зондирования программного комплекса MIRS, который использует микроволновые измерения приборов AMSU/MHS КА серии NOAA, MetOp и прибора ATMS КА Suomi NPP. По результатам анализа ряда данных, не участвовавших в процессе разработки и обучения моде-

лей, за май–октябрь 2019–2020 г. получена точность 84 %.

Дополнительно проведены испытания по 109 метеорологическим станциям зоны ответственности ФГБУ «Западно-Сибирское УГМС», расположенных в границах 49,9–60,35° с.ш. и 75,68–88,67° в.д. За период апрель–сентябрь 2021 г. значение общей оправдываемости свидетельствует о том, что успешно классифицируются 96 % событий (наличие/отсутствие грозы). Нужно отметить, что ложное детектирование гроз происходит при присутствии атмосферных явлениях конвективного характера, например ливневых осадков слабой интенсивности.

Представленный алгоритм динамичен. Процедуру обучения можно инициализировать повторно при накоплении достаточного количества данных. Это позволит учитывать появление новых экстремумов атмосферных характеристик.

Использование представленного алгоритма и результатов его расчетов перспективно в качестве консультативного материала при составлении краткосрочного прогноза погоды в оперативной работе синоптических групп.

Список источников

- Абдуллаев С.М., Ленская О.Ю., Желнин А.А. Структура мезомасштабных конвективных систем в Центральной России // Метеорология и гидрология. 2012. № 1. С. 20–32.
- Алексеева А.А., Бухаров М.В. Диагноз гроз по синхронной информации спутниковых радиометров микроволнового и инфракрасного диапазонов // Метеорология и гидрология. 2005. № 6. С. 29–37.
- Алексеева А.А., Бухаров М.В., Лосев В.М., Соловьев В.И. Диагноз осадков и гроз по измерениям уходящего теплового излучения облачности с геостационарных спутников // Метеорология и гидрология. 2006. № 8. С. 33–42.
- Булыгина О.Н., Веселов В.М., Разуваев В.Н., Александрова Т.М. Описание массива срочных данных об основных метеорологических параметрах на станциях России. Свидетельство о государственной регистрации базы данных № 2014620549. URL: <http://meteo.ru/data/163-basic-parameters#описание-массива-данных>
- Бухаров М.В. Диагноз вероятности гроз по спутниковой информации // Метеорология и гидрология. 2013. № 8. С. 5–16.
- Горбатенко В.П., Кузевская И.В., Пустовалов К.Н., Чурсин В.В., Константинова Д.А. Оценка изменчивости конвективного потенциала атмосферы в условиях изменяющегося климата Западной Сибири // Метеорология и гидрология. 2020. № 5. С. 108–117.
- Джулли А., Пал С. Библиотека Keras – инструмент глубокого обучения. Реализация нейронных сетей с помощью библиотек Theano и TensorFlow. М.: ДМК-Пресс, 2017. 294 с.
- Иванова А.Р. Мировой опыт наукастинга грозовой деятельности // Метеорология и гидрология. 2019. № 11. С. 71–83.
- Орельен Ж. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow. Концепции, инструменты и техники для создания интеллектуальных систем. М.: Вильямс, 2018. 688 с.
- РД 52.27.284-91. Методические указания. Проведение производственных (оперативных) испытаний новых и усовершенствованных методов гидрометеорологических и географических прогнозов. Л.: Гидрометеоиздат, 1991.
- Чернукульский А.В., Козлов Ф.А., Золина О.Г., Булыгина О.Н., Семенов В.А. Климатология осадков разного генезиса в Северной Евразии // Метеорология и гидрология. 2018. № 7. С. 5–18.
- Algorithm Theoretical Basis Document. Microwave Integrated Retrieval System (MIRS) // NOAA. NESDIS (STAR). 2006. 39 p.
- del Moral A., Rigo T., Llasat M.C. A radar-based centroid tracking algorithm for severe weather surveillance: Identifying split/merge processes in convective systems // Atmospheric Research. 2018. V. 213. P. 110–120.
- Galanaki E. et al. Thunderstorm climatology in the Mediterranean using cloud-to-ground lightning observations // Atmospheric Research. 2018. V. 207. P. 136–144.
- Goodman S.J. et al. GLM lightning cluster-filter algorithm, Algorithm Theoretical Basis Document // NOAA NESDIS Center for Satellite Applications and Research. Washington, DC, Ver. 2012. V. 3. P. 73.
- Grassotti C., Liu S., Liu Q., Boukabara S.-A., Garrett K., Iturbide-Sanchez F., Honeyage R. Precipitation Estimation from the Microwave Integrated Retrieval System (MiRS) // Satellite Precipitation Measurement. 2020. V. 1. P. 153–168.
- Hersbach H., Bell B., Berrisford P., Hirahara S., Horányi A., Muñoz-Sabater J., Thépaut J.N. The ERA5 global reanalysis // Quarterly Journal of the Royal Meteorological Society. 2020. V. 146, No. 730. P. 1999–2049.
- Lee J. G. et al. Improvement of the rapid-development thunderstorm (RDT) algorithm for use with the GK2A satellite // Asia-Pacific Journal of Atmospheric Sciences. 2020. V. 56, No. 2. P. 307–319.
- Mecikalski J. R. et al. Cloud-top properties of growing cumulus prior to convective initiation as measured by Meteosat Second Generation. Part I: Infrared fields // Journal of Applied Meteorology and Climatology. 2010. V. 49, No. 3. P. 521–534.
- Meng Q., Yao W., Xu L. Development of lightning nowcasting and warning technique and its application // Advances in meteorology. 2019. V. 2019.
- Network V.L.D. Total Lightning Statistics 2021: Vaisala Annual Lightning Report.
- User Manual. Microwave Integrated Retrieval System (MIRS) // NOAA. NESDIS (STAR). 2016. V. 1.13. 87 p.

References

- Abdullaev S.M., Lenskaya O.Y., Zhelnin A.A. The structure of mesoscale convective systems in Central Russia // Russian Meteorology and Hydrology. 2012. V. 37 (1). P. 12–20.
- Alekseeva A.A., Bukharov M.V. *Diagnoz groz po sinkhronnoy informatsii sputnikovyx radiometrov mikrovolnovogo i infrakrasnogo diapazonov* [Diagnosis of thunderstorms from synchronous information of microwave and infrared satellite radiometers] // Meteorology and Hydrology. 2005. № 6. pp. 29–37. In Russian
- Alekseeva A.A., Bukharov M.V., Losev V.M., Solov'ev V.I. *Diagnoz osadkov i groz po izmereniyam ukhodyashchego teplovogo izlucheniya oblachnosti s geostatsionarnykh sputnikov* [Forecasting precipitation and thunderstorms from geostationary satellite measurements of outgoing thermal radiation from clouds] // Meteorology and Hydrology. 2006. № 8. pp. 33–42. In Russian
- Bulygina O.N., Veselov V.M., Razuvaev V.N., Aleksandrova T.M. *Opisanie massiva srochnykh dannykh ob osnovnykh meteorologicheskikh parametrah na stantsiyakh Rossii. Svidetel'stvo o gosudarstvennoy registratsii bazy dannykh № 2014620549* [Description of the array of urgent data on the main meteorological parameters at the stations of Russia. Certificate of state registration of the database № 2014620549] // <http://meteo.ru/data/163-basic-parameters#описание-массива-данных>. In Russian
- Bukharov M.V. Satellite diagnosis of thunderstorm probability // Russian Meteorology and Hydrology. 2013. V. 38 (8). pp. 515–521.
- Gorbatenko V.P., Kuzhevskaya I.V., Pustovalov K.N., Chursin V.V., Konstantinova D.A. Assessment of Atmospheric Convective Potential Variability in Western Siberia in Changing Climate // Russian Meteorology and Hydrology. 2020. V. 45 (5). pp. 360–367.
- Gulli A., Pal S. Deep learning with Keras // Packt Publishing Ltd. 2017. 286 p.
- Ivanova A.R. International Practices of Thunderstorm Nowcasting // Russian Meteorology and Hydrology. 2019. V. 44 (11). pp. 756–763.
- Aurelien G. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems // O'Reilly Media. 2017. 856 p.
- RD 52.27.284-91. *Metodicheskie ukazaniya. Provedenie proizvodstvennykh (operativnykh) ispytaniy novykh i usovershenstvovannykh metodov gidrometeorologicheskikh i geliograficheskikh prognozov.* [Methodological guidelines. Conducting production (opera-

tional) tests of new and improved methods of hydrometeorological and heliographic forecasts.] // Leningrad: Gidrometeoizd. 1991. In Russian

Chernokulsky A.V., Kozlov F.A., Semenov V.A., Zolina O.G., Bulygina O.N. Climatology of precipitation of different genesis in Northern Eurasia // Russian Meteorology and Hydrology. 2018. V. 43 (7). pp. 425–435.

Algorithm Theoretical Basis Document. Microwave Integrated Retrieval System (MIRS) // NOAA. NESDIS (STAR). 2006. 39 p.

del Moral A., Rigo T., Llasat M.C. A radar-based centroid tracking algorithm for severe weather surveillance: Identifying split/merge processes in convective systems // Atmospheric Research. 2018. V. 213. pp. 110–120.

Galanaki E. et al. Thunderstorm climatology in the Mediterranean using cloud-to-ground lightning observations // Atmospheric Research. 2018. V. 207. pp. 136–144.

Goodman S. J. et al. GLM lightning cluster-filter algorithm, Algorithm Theoretical Basis Document // NOAA NESDIS Center for Satellite Applications and Research, Washington, DC, Ver. 2012. V. 3. 73 p.

Grassotti C., Liu S., Liu Q., Boukabara S.-A., Garrett K., Iturbide-Sanchez F., Honeyage R. Precipitation Estimation from the Microwave Integrated Retrieval System (MiRS) // Satellite Precipitation Measurement. 2020. V. 1. pp. 153–168.

Hersbach H., Bell B., Berrisford P., Hirahara S., Horányi A., Muñoz-Sabater J., Thépaut J. N. The ERA5 global reanalysis // Quarterly Journal of the Royal Meteorological Society. 2020. V. 146 (730). pp. 1999–2049.

Lee J. G. et al. Improvement of the rapid-development thunderstorm (RDT) algorithm for use with the GK2A satellite // Asia-Pacific Journal of Atmospheric Sciences. 2020. V. 56 (2). pp. 307–319.

Mecikalski J. R. et al. Cloud-top properties of growing cumulus prior to convective initiation as measured by Meteosat Second Generation. Part I: Infrared fields // Journal of Applied Meteorology and Climatology. 2010. V. 49 (3). pp. 521–534.

Meng Q., Yao W., Xu L. Development of lightning nowcasting and warning technique and its application // Advances in meteorology. 2019. V. 2019.

Network V. L. D. Total Lightning Statistics 2021: Vaisala Annual Lightning Report. 2021.

User Manual. Microwave Integrated Retrieval System (MIRS) // NOAA. NESDIS (STAR). 2016. V. 1.13. 87 p.

Информация об авторах:

Чурсин В.В., аспирант, младший научный сотрудник, кафедра метеорологии и климатологии, ПНИЛГК, геолого-географический факультет, Национальный исследовательский Томский государственный университет, Томск, Россия.
E-mail: skriptym@mail.ru

Кужевская И.В., кандидат географических наук, доцент, кафедра метеорологии и климатологии, геолого-географический факультет, Национальный исследовательский Томский государственный университет, Томск, Россия.
E-mail: irina-kuz@vtomske.ru

Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Chursin V.V., Postgraduate Student, Department of Meteorology and Climatology, Geology and Geography Faculty, National Research Tomsk State University, Tomsk, Russia.
E-mail: skriptym@mail.ru

Kuzhevskaya I.V., Cand. Sci. (Geography), Associate Professor, Department of Meteorology and Climatology, Geology and Geography Faculty, National Research Tomsk State University, Tomsk, Russia.
E-mail: irina-kuz@vtomske.ru

The authors declare no conflicts of interests.

Статья поступила в редакцию 03.05.2022; одобрена после рецензирования 08.08.2022; принята к публикации 12.09.2022

The article was submitted 03.05.2022; approved after reviewing 08.08.2022; accepted for publication 12.09.2022